ANSWER KEY Makeup Exam:

Test Form A
1. D
2. D
3. C
4. A
5. A
6. B
7. C
8. A
9. A
10. E
11. E
12. E
13. D
14. A
15. C
16. B
17. C
18. D
19. E
20. B
21. B
22. E
23. B
24. B
25. A
26. D
27. C
28. D
29. D
30. B
31. D
32. D

Test Form E
1. E
2. C
3. C
4. D
5. A
6. E
7. C
8. B
9. A
10. C
11. E
12. D
13. D
14. D
15. C
16. A
17. B
18. A
19. C
20. E
21. A
22. D
23. A
24. B
25. B
26. D
27. B
28. D
29. A
30. E
31. A
32. A

Solutions for Form E appear next — you can find the corresponding questions on Form A to see how to do the problems.
Questions 1 – 5 A study conducted by the General Social Survey asked a random sample of American adults how they would classify their family income and their happiness level. The data is summarized below.

<table>
<thead>
<tr>
<th>Family Income</th>
<th>Not too Happy</th>
<th>Pretty Happy</th>
<th>Very Happy</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Above Average</td>
<td>26</td>
<td>233</td>
<td>(164)</td>
<td>423</td>
</tr>
<tr>
<td>Average</td>
<td>117</td>
<td>473</td>
<td>293</td>
<td>883</td>
</tr>
<tr>
<td>Below Average</td>
<td>172</td>
<td>383</td>
<td>132</td>
<td>687</td>
</tr>
<tr>
<td>Total</td>
<td>315</td>
<td>1089</td>
<td>589</td>
<td>1993</td>
</tr>
</tbody>
</table>

1. What is the contribution to the test statistic of the Above Average Income / Very Happy category?
   a) 164  b) 39  c) 9.27  d) 125.01  
   ![Image showing computation](image)

2. The $p$-value for this test was reported by the statistical computing package as 0.000. So we can conclude:
   a) income determines happiness in the sample.
   b) happiness determines income in the population.
   c) income is associated with happiness in the population.
   d) income is associated with happiness in the sample.
   e) income is not associated with happiness in the sample.

For the rest of this problem we will concentrate on people who consider themselves Pretty Happy. In particular, we want to compare the Above Average and the Below Average Income groups.

3. The point estimators of the proportion of Americans in the Above Average Income group that consider themselves Pretty Happy and the proportion of Americans in the Below Average Income group that consider themselves Pretty Happy are, respectively:
   a) .21 and .35  b) .12 and .19  c) .55 and .56  d) .49 and .81

4. The 95% confidence interval to compare $\pi_1$ and $\pi_2$ is (-.067, 0.0536). So we are 95% confident that the proportion of Americans with Above Average Income that consider themselves Pretty Happy is than the proportion of Americans with Below Average Income that consider themselves the same.
   a) significantly higher  b) significantly lower  c) significantly different  d) not significantly different

5. Suppose we use this data to test $H_0: \pi_1 - \pi_2 = 0$ vs. $H_A: \pi_1 - \pi_2 < 0$. What can we say about the $p$-value?
   a) it is greater than .05  b) it is less than .05  c) it is less than .025  d) cannot say anything about the $p$-value

   ![Image showing computation](image)
Questions 6-7: Studies have shown that the presence of music can alter people's perceptions of their other senses, including their sense of taste. In particular, restaurants have found that guests may consistently report that their food tasted better or worse depending on the characteristics of the music played while they were dining. The following plot represents the average taste of a particular menu item (on a scale from 1-10 as judged by diners) in the presence of music played at three different volumes (soft, moderate, and loud) from two different sources (live or recorded).

6. The two factors in this study are
a) recorded and taste of food.
b) soft and music source.
c) live and recorded.
d) music source and recorded.
( ) music volume and music source.

7. Which of the following interpretations of this graph is most clearly correct?
a) In general, food tastes better with soft music than with other volumes, from any source. ☒
b) In general, food tastes better with loud music than with other volumes, from any source. ☒
c) In general, food tastes better with live music than with recorded music, at any volume. ☒
d) In general, food tastes better with recorded music than with live music, at any volume. ☒
e) Whether food tastes better with live or recorded music depends on the volume at which the music is played. ☒

Questions 8 - 12: It is not surprising that seemingly irrelevant characteristics such as physical attractiveness can subconsciously influence employers' decisions in the hiring process. However, some people believe that such traits may actually help employers predict future job performance in a quantifiable way. Each of the five situations below describes an inference that we would like to make about a different parameter. Match each of the five situations below with the parameter of interest from the list.

8) one mean               9) one proportion
7) difference of two independent means
6) difference of two independent proportions
5) mean of matched paired differences

8. This belief is widespread but not the majority opinion: Only 37.5% of managers said that they believed physical attractiveness and job performance could be connected in a scientific way.

9. People may have an inflated view of their own looks: The participants in the study were each asked to rate their own attractiveness on a scale from 1-10, and the average answer was 6.6.

10. The people rated as "attractive" may be less efficient than the people rated as "average" or "unattractive", but perhaps only slightly: The "attractive" group completed only 22.4 tasks per week on average, compared to 25.3 for the other group.

11. People rated as "attractive" and "unattractive" were partnered off and asked to each complete the same task while competing against each other to finish first. The people rated as "attractive" finished in 93.9 minutes on average, while the people rated as "unattractive" took 95.1 minutes.

12. The ability to learn new skills lacks any strong relationship to attractiveness: 66.2% of people rated as "attractive" passed a competency evaluation at the end of training, compared to 63.4% of people rated as "average" or "unattractive."
Questions 13 – 19: A variety of things can affect the productivity of office workers, including the amount of space each person is given in which to work. A manager at a certain corporation wants to determine which types of office arrangements allow workers to get more done, on average. He collects data on employees whose job is to complete the same task. The following is a partial computer output for comparing the number of minutes taken to complete the task by employees working in four different types of office environments.

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
<th>P</th>
<th>Level</th>
<th>N</th>
<th>Mean</th>
<th>StDev</th>
</tr>
</thead>
<tbody>
<tr>
<td>Factor</td>
<td>3</td>
<td>18642</td>
<td>6214.0</td>
<td>4.475</td>
<td>0.018</td>
<td>Own Offices</td>
<td>5</td>
<td>96.8</td>
<td>20.86</td>
</tr>
<tr>
<td>Error</td>
<td>16</td>
<td>22217</td>
<td>1388.5</td>
<td></td>
<td></td>
<td>Desks, Open</td>
<td>5</td>
<td>109.2</td>
<td>34.24</td>
</tr>
<tr>
<td>Total</td>
<td>19</td>
<td>40859</td>
<td></td>
<td></td>
<td></td>
<td>Cubicles</td>
<td>5</td>
<td>126.6</td>
<td>34.23</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Outdoors</td>
<td>5</td>
<td>177.0</td>
<td>52.68</td>
</tr>
</tbody>
</table>

13. How many replications are used in this study?  
   a) 1  b) 3  c) 4  d) 5  e) 20

14. Which of these is a correct interpretation of the null hypothesis here, in words?  
   a) All four environments have the same number of workers, on average.  
   b) All four environments do not have the same number of workers, on average.  
   c) Workers do not take the same amount of time to complete a task, on average, in all four environments.  
   d) Workers take the same amount of time to complete a task, on average, in all four environments.

15. At a significance level of 0.05, we can conclude from the ANOVA test that the population means of the four groups are:  
   a) plausibly all equal.  b) all different.  c) not all equal.  d) equal to the population variances of the corresponding groups.

16. When making all pairwise comparisons for this problem with confidence level of 94% for EACH interval, what would be the family confidence level?  
   a) 64%  b) 88%  c) 94%  d) 99%  e) 100%

17. The pooled standard deviation is:  
   a) 21.5  b) 37.3  c) 64.6  d) 111.9  e) 35.50

18. When making all pairwise comparisons with 95% family confidence, the margin of error was determined to be 76.6. Which of the following is the best summary of the results?  
   a) OO 96.8  b) OO 96.8  c) OO 96.8  d) OO 96.8  e) Connect any pair of means that are less than 76.6 minutes apart.

19. If our study also divided the workers in each environment into a group who worked with background music and a group who worked in silence, then we would be adding to the study:  
   a) one more level for each of two existing factors.  b) two more levels for an existing factor.  c) one more factor, with two levels.  d) two more factors, with one level each.  e) none of the above—the numbers of factors and levels would not change.

20. We make statistical inferences when we know:  
   a) the population size but not the data for the whole population  b) the sample size but not the population size  c) the sample size but not the data for the whole sample  d) the population size but not the sample size  e) the data for the whole sample but not for the whole population

21. We use the FPC in statistical inference when we know:  
   a) the population size but not the data for the whole population  b) the sample size but not the population size  c) the sample size but not the data for the whole sample  d) the population size but not the sample size  e) the data for the whole sample but not for the whole population
Questions 22–28 The National Math and Science Initiative (NMSI) has recently begun a controversial program in which high school students are paid cash incentives for passing an end-of-year standardized test. Suppose we conduct a similar study, in which end-of-year test scores (y) are measured on a scale of 0–100 and the amount of the cash incentive offered to the student (x) is measured in dollars from $0 to $500. A scatterplot of the 96 observations in the sample and the regression line is shown below, along with parts of the computer output (with some information intentionally left blank).

<table>
<thead>
<tr>
<th>Coeff</th>
<th>Std Error</th>
<th>t Stat</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept 67.51</td>
<td>2.509</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cash 0.0148</td>
<td>0.00886</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>X Value</td>
</tr>
<tr>
<td>Confidence Level</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>For Average Y</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interval Half Width</td>
</tr>
<tr>
<td>Confidence Interval Lower Limit</td>
</tr>
<tr>
<td>Confidence Interval Upper Limit</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>For Individual Response Y</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interval Half Width</td>
</tr>
<tr>
<td>Prediction Interval Lower Limit</td>
</tr>
<tr>
<td>Prediction Interval Upper Limit</td>
</tr>
</tbody>
</table>

22. The coefficient 0.01476 in the regression equation is  
   a) the parameter $\beta_0$.  
   b) the parameter $\beta_1$.  
   c) our estimate of the parameter $\beta_0$.  
   d) our estimate of the parameter $\beta_1$.  

23. Which of the following is the best interpretation of the slope? 
   a) For each additional dollar offered to students, their scores increase by 0.01476 points, on average.  
   b) For each additional dollar offered to students, their scores increase by 67.51 points, on average.  
   c) For each 0.01476 additional dollars offered to students, their scores increase by one point, on average.  
   d) For each 67.51 additional dollars offered to students, their scores increase by one point, on average. 

24. Which of the following is the best interpretation of the intercept?  
   a) The predicted score of a student who is offered no cash is 0.01476 points.  
   b) The predicted score of a student who is offered no cash is 67.51 points.  
   c) The amount of cash offered to a student who scores a zero is, on average, 0.01476 dollars.  
   d) The amount of cash offered to a student who scores a zero is, on average, 67.51 dollars.  
   e) None of the above; it is not appropriate to interpret the intercept in this situation. 

25. Calculate the predicted test score of a student who is offered a cash incentive of $200.  
   $\hat{y} = 67.51 + 0.01476(200)$  
   a) 64.56  
   b) 70.47  
   c) 73.41  
   d) 76.37  
   e) 82.27 

26. The p-value for the ANOVA test was 0.0952, so there is ________ evidence that scores on the test depend on the size of the cash incentive.  
   a) not enough  
   b) pretty strong  
   c) very strong  
   d) some  
   e) no 

27. Calculate the value of the $t$ test statistic for testing whether score depends on cash incentive.  
   $t = \frac{\text{Coeff} - 0}{\text{StdErr}} = \frac{0.01476}{0.00886}$  
   a) 0.600  
   b) 1.67  
   c) 2.79  
   d) 4.10  
   e) 0.599 

28. Which of the four intervals labeled as [A]–[D] in the Minitab output would be the widest?  
   a) [A]  
   b) [B]  
   c) [C]  
   d) [D]  
   e) All four would have the same width.  

   $\Rightarrow$ PI wider than CI  
   $\Rightarrow$ Intervals are wider when $X$ far from $\bar{X}$  

   $\Rightarrow$ PI when $X=500$
Questions 29 – 32  One important feature of a digital camera is the battery life, measured as the number of shots taken until the battery needs to be recharged. The following data set contains the battery life of 29 subcompact cameras and 16 compact cameras as reported by Consumer Reports, July 2009.

<table>
<thead>
<tr>
<th>Subcompact</th>
<th>Compact</th>
</tr>
</thead>
<tbody>
<tr>
<td>140</td>
<td>130</td>
</tr>
<tr>
<td>150</td>
<td>150</td>
</tr>
<tr>
<td>160</td>
<td>200</td>
</tr>
<tr>
<td>170</td>
<td>220</td>
</tr>
<tr>
<td>200</td>
<td>230</td>
</tr>
<tr>
<td>200</td>
<td>240</td>
</tr>
<tr>
<td>200</td>
<td>260</td>
</tr>
<tr>
<td>200</td>
<td>260</td>
</tr>
<tr>
<td>200</td>
<td>260</td>
</tr>
<tr>
<td>205</td>
<td>280</td>
</tr>
<tr>
<td>210</td>
<td>330</td>
</tr>
<tr>
<td>210</td>
<td>350</td>
</tr>
<tr>
<td>220</td>
<td>400</td>
</tr>
<tr>
<td>220</td>
<td>400</td>
</tr>
<tr>
<td>220</td>
<td>450</td>
</tr>
<tr>
<td>230</td>
<td>450</td>
</tr>
<tr>
<td>240</td>
<td>500</td>
</tr>
</tbody>
</table>

Summary statistics:

<table>
<thead>
<tr>
<th>Column</th>
<th>n</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Median</th>
<th>Q1</th>
<th>Q3</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>subcompact</td>
<td>29</td>
<td>234.65517</td>
<td>49.20676</td>
<td>230</td>
<td>200</td>
<td>280</td>
<td>140</td>
<td>330</td>
</tr>
<tr>
<td>compact</td>
<td>16</td>
<td>303.125</td>
<td>111.99516</td>
<td>270</td>
<td>225</td>
<td>400</td>
<td>130</td>
<td>500</td>
</tr>
</tbody>
</table>

29. Which of the following procedures would NOT be appropriate to use to analyze this data?
   I   test for the mean of dependent samples
   II  test for two means, variances unknown but assumed equal
   III test for two means, variances unknown and separate
   IV  Wilcoxon Rank Sum test

   a) I and II         b) II and III       c) II and IV       d) III and IV       e) I and IV

30. The Wilcoxon Rank Sum Test gives a two-sided p-value of .05445. Assuming this test is appropriate for the data, the conclusions would be that at α=0.05:
   a) the mean battery life for subcompact cameras is significantly different than for compact cameras.
   b) the median battery life for subcompact cameras is significantly different than for compact cameras.
   c) the std. deviation of battery life for subcompact cameras is significantly different than for compact cameras.
   d) the mean battery life for subcompact cameras is not significantly different than for compact cameras.
   e) the median battery life for subcompact cameras is not significantly different than for compact cameras.

31. The Pooled Variance T Test gives a two-sided p-value of .0067. Assuming this test is appropriate for the data, the conclusions would be that at α=0.05:
   a) the mean battery life for subcompact cameras is significantly different than for compact cameras.
   b) the median battery life for subcompact cameras is significantly different than for compact cameras.
   c) the std. deviation of battery life for subcompact cameras is significantly different than for compact cameras.
   d) the mean battery life for subcompact cameras is not significantly different than for compact cameras.
   e) the median battery life for subcompact cameras is not significantly different than for compact cameras.

32. Construct an approximately 95% CI for the median battery life of subcompact digital cameras.
   a) (210, 260)     b) (200, 280)     c) (200, 450)     d) (225, 400)     e) (205, 280)

\[ 4n - 2 = \text{A}(29) - 2 = 9.6 \]
\[ \text{Round to 10} \]
\[ \text{Find obs in 10th pos from bottom} + \text{to top} \]